A new transform for solving the noisy complex exponentials approximation problem

P. Barone

Istituto per le Applicazioni del Calcolo "M. Picone", C.N.R.,

Viale del Policlinico 137, 00161 Rome, Italy

e-mail: barone@iac.rm.cnr.it

fax: 39-6-4404306

Abstract

The problem of estimating a complex measure made up by a linear combination of Dirac distributions centered on points of the complex plane from a finite number of its complex moments affected by additive i.i.d. Gaussian noise is considered. A random measure is defined whose expectation approximates the unknown measure under suitable conditions. An estimator of the approximating measure is then proposed as well as a new discrete transform of the noisy moments that allows to compute an estimate of the unknown measure. A small simulation study is also performed to experimentally check the goodness of the approximations.

Key words and phrases: Complex moments; Pade' approximants; logarithmic potentials; random determinants; random polynomials; pencils of matrices

Introduction

Let us consider the complex measure defined on a compact set $D \subset \mathbb{C}$ by

$$S(z) = \sum_{j=1}^{p} c_j \delta(z - \xi_j), \quad \xi_j \in \mathrm{int}(D), \quad c_j \in \mathbb{C}$$

and let be

$$s_k = \int_D z^k S(z) dz = \iint_D (x+iy)^k S(x+iy) dx dy, \quad k = 0, 1, 2, \ldots$$

the complex moments. It turns out that

$$s_k = \sum_{j=1}^{p} c_j \xi_j^k. \tag{1}$$

Let us assume to know an even number $n \geq 2p$ of noisy complex moments

$$\mathbf{a}_k = s_k + \boldsymbol{\nu}_k, \quad k = 0, 1, 2, \ldots, n-1$$

where $\boldsymbol{\nu}_k$ is a complex Gaussian, zero mean, white noise, with finite known variance σ^2. In the following all random quantities are denoted by bold characters. We want to estimate $S(z)$ from $\{\mathbf{a}_k\}_{k=0,\ldots,n-1}$. From equation (1) this is equivalent to estimate $p, c_j, \xi_j, \ j = 1, \ldots, p$, which is the well known difficult problem of complex exponentials approximation.

The problem is central in many disciplines and appears in the literature in different forms and contexts (see e.g. [6,12,22,24,28]). The assumptions about the noise variance (constant and known) are made here to simplify the analysis. However in many applications the noise is an instrumental one which is well represented by a white noise, zero mean, Gaussian process whose variance is known or easy to estimate. A

typical example is provided by NMR spectroscopy (see e.g. [8]).

In the noiseless case the problem becomes the complex exponential interpolation problem [14]. Conditions for existence and unicity of the solution are ([14, Th.7.2c]):

$$det U_0(\underline{s}) \neq 0, \quad det U_1(\underline{s}) \neq 0$$

where

$$U(s_0, \ldots, s_{2p-2}) = \begin{bmatrix} s_0 & s_1 \ldots s_{p-1} \\ s_1 & s_2 \ldots s_p \\ . & . \quad \ldots . \\ s_{p-1} & s_p \ldots s_{2p-2} \end{bmatrix}$$

and

$$U_0(\underline{s}) = U(s_0, \ldots, s_{2p-2}), \quad U_1(\underline{s}) = U(s_1, \ldots, s_{2p-1}).$$

In fact exactly $n = 2p$ noiseless moments are sufficient to fully retrieve $S(z)$, where

$$p = \max_{n \in I\!N} \{n \mid det(U(s_0, \ldots, s_{n-2})) \neq 0\}.$$

Moreover $(\xi_j, \ \ j = 1, \ldots, p)$ are the generalized eigenvalues of the pencil $P = [U_1(\underline{s}), U_0(\underline{s})]$ i.e. they are the roots of the polynomial in the variable z

$$det[U_1(\underline{s}) - zU_0(\underline{s})]$$

and c_j are related to the generalized eigenvector $\underline{u}_j$ of P by $c_j = \underline{u}_j^T [s_0, \ldots, s_{p-1}]^T$. In fact from equation (1) we have $\underline{c} = V^{-1}[s_0, \ldots, s_{p-1}]^T$ where

$$V = Vander(\xi_1, \ldots, \xi_p)$$

is the square Vandermonde matrix based on $(\xi_1, \ldots, \xi_p)$. But it easy to show (see e.g. [2]) that

$$U_0(\underline{s}) = VCV^T, \;\; U_1(\underline{s}) = VCZV^T$$

where

$$C = diag\{c_1, \ldots, c_p\} \text{ and } Z = diag\{\xi_1, \ldots, \xi_p\}.$$

Therefore $\underline{u}_k = V^{-T}\underline{e}_k$ is the right generalized eigenvector of P corresponding to ξ_k, where $\underline{e}_k$ is the $k-$th column of the identity matrix I_p of order p.

Viceversa when $s_k = 0, \; \forall k$ it was proved in [15] that

$$det[U(\mathbf{a}_0, \ldots, \mathbf{a}_{n-2})] = det[U_0(\underline{\mathbf{a}})] \neq 0 \;\; \forall n \text{ a.s.}$$

and

$$det[U(\mathbf{a}_1, \ldots, \mathbf{a}_{n-1})] = det[U_1(\underline{\mathbf{a}})] \neq 0 \;\; \forall n \text{ a.s..}$$

Moreover associated to the random polynomial

$$det[U_1(\underline{\mathbf{a}}) - zU_0(\underline{\mathbf{a}})] \tag{2}$$

a condensed density $h_n(z)$ can be considered which is the expected value of the (random) normalized counting measure on the zeros of this polynomial i.e.

$$h_n(z) = \frac{2}{n} E\left[\sum_{j=1}^{n/2} \delta(z - \boldsymbol{\xi}_j)\right].$$

It was proved in [1] that if $z = re^{i\theta}$, the marginal condensed density $h_n^{(r)}(r)$ w.r. to r of the generalized eigenvalues is asymptotically in n a Dirac δ supported on the unit circle $\forall\sigma^2$. Moreover for finite n the the marginal condensed density w.r. to θ is uniformly distributed on $[-\pi, \pi]$. Starting from the generalized eigenvalues $\boldsymbol{\xi}_j$ and

generalized eigenvectors $\underline{\mathbf{u}}_j$ of the pencil

$$\mathbf{P} = [U(\mathbf{a}_1, \ldots, \mathbf{a}_{n-1}), U(\mathbf{a}_0, \ldots, \mathbf{a}_{n-2})]$$

we then define a family of random measures

$$\mathbf{S}_n(z) = \sum_{j=1}^{n/2} \mathbf{c}_j \delta(z - \boldsymbol{\xi}_j)$$

where $\mathbf{c}_j = \underline{\mathbf{u}}_j^T [\mathbf{a}_0, \ldots, \mathbf{a}_{n/2-1}]^T$ and we give conditions under which $E[\mathbf{S}_n(z)]$ approximates $S(z)$. Moreover we define a discrete transform (P-Transform) on a lattice of points on D, which is an unbiased and consistent estimator of $E[\mathbf{S}_n(z)]$ on the lattice thus providing a computational device to solve the original problem.

In [4] the same problem was afforded. The joint distribution of the coefficients of the random polynomial (2) (when $s_k \neq 0, \ \forall k$) was approximated by a multivariate Gaussian distribution and a theorem by Hammersley [7] was used to compute the associated condensed density of its roots. An heuristic algorithm was then used to identify the main peaks of the condensed density and to get estimates of $p, \ \xi_j$ and $c_j, \ j = 1, \ldots, p$ based on them. In the present work the ideas presented in [4] are put on a more rigorous mathematical framework. A different approximation of the condensed density is considered and an automatic estimation procedure is proposed.

The paper is organized as follows. In the first section we study the distribution of the generalized eigenvalues of the random pencil $\mathbf{P}$ and we give an easily computable approximate expression of the associated condensed density. In section 2 we consider the identifiability problem for $S(z)$ given the data $\underline{\mathbf{a}}$. Conditions for identifiability are given and the approximation properties of $E[\mathbf{S}_n(z)]$ are proved. In section 3 the

P-transform is defined and its statistical properties are studied. In section 4 the procedure for estimating the parameters p, $\{\xi_j, c_j, j = 1, \ldots, p\}$ of the unknown measure from the P-transform is described. Finally in section 5 some experimental results on synthetic data are reported.

1 Distribution of the generalized eigenvalues of the pencil P

We start by making some technical assumptions on the noise model. When $s_k = 0\ \forall k$, we noticed in the introduction that ξ_j are, asymptotically on n, uniformly distributed on the unit circle. Therefore, when $s_k \neq 0$ is given by (1), we can assume that $n_p = n/2 - p$ among the $\xi_j, j = 1, \ldots, n/2$ are related to noise and then they can be modeled for large n by $\tilde{\xi}_j = e^{\frac{2\pi i j}{n_p}}$ i.e. by uniformly spaced deterministic generalized eigenvalues. Therefore the Vandermonde matrix based on $\tilde{\xi}_j, j = 1, \ldots, n_p$ is simply given by $V = \sqrt{n_p} \cdot F \in \mathbb{C}^{n_p \times n_p}$ where $F_{hk} = \frac{1}{\sqrt{n_p}} e^{\frac{2\pi i h k}{n_p}}$ is the discrete Fourier transform matrix. Hence

$$\underline{\tilde{\mathbf{c}}} = V^{-1}[\boldsymbol{\nu}_0, \ldots, \boldsymbol{\nu}_{n_p-1}]^T = \frac{1}{\sqrt{n_p}} F^H [\boldsymbol{\nu}_0, \ldots, \boldsymbol{\nu}_{n_p-1}]^T$$

and $\underline{\tilde{\mathbf{c}}}$ has a complex multivariate Gaussian distribution with

$$E[\tilde{\mathbf{c}}_j] = 0 \text{ and } E[\tilde{\mathbf{c}}_j \overline{\tilde{\mathbf{c}}_h}] = \frac{\sigma^2}{n_p} \delta_{jh}.$$

Based on these observations we define a new noise process as

$$\tilde{\boldsymbol{\nu}}_k = \begin{cases} \sum_{j=1}^{n_p} \tilde{\mathbf{c}}_j \tilde{\xi}_j^k, & k < n_p \\ \boldsymbol{\nu}_k, & k \geq n_p \end{cases}$$

and we assume that $\underline{\tilde{\mathbf{c}}}$ is independent of $\boldsymbol{\nu}_k, k \geq n_p$. But then $E[\tilde{\boldsymbol{\nu}}_k] = 0$ and

$$E[\tilde{\boldsymbol{\nu}}_k \overline{\tilde{\boldsymbol{\nu}}}_h] = \begin{cases} \sum_{i,j}^{1,n_p} \tilde{\xi}_i^k \overline{\tilde{\xi}}_j^h E[\tilde{\mathbf{c}}_i \overline{\tilde{\mathbf{c}}}_j] = \frac{\sigma^2}{n_p} \sum_{r=1}^{n_p} e^{\frac{2\pi i r(k-h)}{n_p}} = \sigma^2 \delta_{hk}, \; k,h < n_p \\ \sum_{j=1}^{n_p} E[\tilde{\mathbf{c}}_j \overline{\boldsymbol{\nu}}_h] \tilde{\xi}_j^k = 0, \qquad h \geq n_p, k < n_p \\ E[\boldsymbol{\nu}_k \overline{\boldsymbol{\nu}}_h] = \sigma^2 \delta_{hk}, \qquad h, k \geq n_p \end{cases}$$

We have then proved the following

Lemma 1 *The random vectors $\boldsymbol{\nu}_k$ and $\tilde{\boldsymbol{\nu}}_k$, $\;k = 0, \ldots, n-1$ are equal in distribution.*

As a consequence in the following we will use $\tilde{\boldsymbol{\nu}}_k$ without loss of generality.

Remark 1 *We notice that when $s_k \neq 0$, if the signal-to-noise ratio is defined as $SNR = \frac{1}{\sigma} \min_{h=1,p} |c_h|$ we have*

$$E[|\tilde{\mathbf{c}}_j|^2] = \frac{\sigma^2}{n_p} = \frac{\min_{h=1,p} |c_h|^2}{n_p SNR^2}.$$

If $SNR \gg \sqrt{\frac{1}{n_p}}$ then $E[|\tilde{\mathbf{c}}_j|^2] \ll |c_k|^2, \;\forall j, k.$

A basic result which will be used extensively in the following is given by

Lemma 2 *Let $T = (T^{(1)}, T^{(2)})$ be the transformation that maps every realization $\underline{a}(\emptyset)$ of $\underline{\mathbf{a}}$ to $(\underline{\xi}(\emptyset), \underline{c}(\emptyset))$ given by $a_k(\emptyset) = \sum_{j=1}^{n/2} c_j(\emptyset) \xi_j(\emptyset)^k, \;\; k = 0, \ldots, n-1,$ where $\emptyset \in \Omega$ and Ω is the space of events. Then $\mathbf{T}$ is a.s. one-to-one. Moreover, for $\sigma \to 0$*

and for $j = 1, \ldots, n/2$

$$E[\boldsymbol{\xi}_j] = \begin{cases} \xi_j + o(\sigma) & j = 1, \ldots, p \\ \\ \tilde{\xi}_{j-p} + o(\sigma), & j = p+1, \ldots, n/2 \end{cases}$$

$$E[\mathbf{c}_j] = \begin{cases} c_j + o(\sigma), \; j = 1, \ldots, p \\ \\ o(\sigma), \qquad\quad j = p+1, \ldots, n/2 \end{cases}$$

<u>proof</u>

From [15] we know that a.s. $det[U_h(\underline{\boldsymbol{\nu}})] \neq 0, \; h = 0, 1$. Moreover, with probability 1, there is no functional dependence between $\underline{\boldsymbol{\nu}}$ and $\underline{s}$. Therefore a.s. $det[U_h(\underline{\mathbf{a}})] \neq 0, \; h = 0, 1$. But then a.s. the complex exponential interpolation problem for $\underline{\mathbf{a}}$ has an unique solution $\forall \emptyset$ hence $\mathbf{T}$ is a.s. one-to-one. The second part of the thesis is based on a Taylor expansion of $\mathbf{T}$ around a suitable point $\underline{x}_0$. A natural candidate for $\underline{x}_0$ would be $\underline{s}$. However we notice that $T^{(1)}(\underline{s})$ is not defined if $n > 2p$, and, as a consequence, also $T^{(2)}(\underline{s})$ is not defined in this case. Therefore, by using Lemma 1, without loss of generality, we assume that the noise is represented by $\tilde{\boldsymbol{\nu}}_k$ i.e.

$$\mathbf{a}_k = \begin{cases} \sum_{j=1}^{p} c_j \xi_j^k + \sum_{j=p+1}^{n/2} \tilde{\mathbf{c}}_{j-p} \tilde{\xi}_{j-p}^k, & k = 0, \ldots, n_p - 1 \\ \\ \sum_{j=1}^{p} c_j \xi_j^k + \boldsymbol{\nu}_k, & k = n_p, \ldots, n-1 \end{cases}$$

where $n_p = n/2 - p$. We then define a new sequence $\tilde{s}_k$ by

$$\tilde{s}_k = \sum_{j=1}^{p} c_j \xi_j^k + \sigma^\alpha \sum_{j=p+1}^{n/2} \tilde{\xi}_{j-p}^k, \quad \alpha \geq 2, \quad k = 0, \ldots, n-1$$

and we consider the process $\mathbf{a}_k$ as a perturbation of $\tilde{s}_k$. Therefore we choose $\underline{x}_0 = \underline{\tilde{s}}$ and notice that

$$T^{(1)}(\underline{\tilde{s}})_j = \begin{cases} \xi_j & j = 1, \ldots, p \\ \tilde{\xi}_{j-p}, & j = p+1, \ldots, n/2 \end{cases}$$

$$T^{(2)}(\underline{\tilde{s}})_j = \begin{cases} c_j & j = 1, \ldots, p \\ \sigma^\alpha, & j = p+1, \ldots, n/2 \end{cases}$$

We now prove that each component of $T^{(1)}(\underline{a})$ is an analytic function of $\underline{a}$ when $\underline{a}$ belong to small neighbor of $\underline{\tilde{s}}$. The proof follows closely [27][Th.6.9.8]. For each fixed ø, the polynomial

$$\phi(z, \underline{a}) = det[U_1(\underline{a}) - zU_0(\underline{a})]$$

is an analytic function of z and $\underline{a}$. Let $\hat{\xi}$ be a zero of $\phi(z, \underline{\tilde{s}})$ and

$$K = \{\zeta || \zeta - \hat{\xi}| = r\}, \quad r > 0$$

be a circle around $\hat{\xi}$ not containing any other generalized eigenvalue of the pencil

$$\tilde{P} = [U(\tilde{s}_1, \ldots, \tilde{s}_{n-1}), U(\tilde{s}_0, \ldots, \tilde{s}_{n-2})].$$

We want to show that K does not pass through any zero of $\phi(z, \underline{a})$. In fact by the definition of K it follows that

$$\inf_{\zeta \in K} |\phi(\zeta, \underline{\tilde{s}})| > 0.$$

But $\phi(z, \underline{a})$ depends continuously on $\underline{a}$, hence there exists $B = \{\underline{x} \in \mathbb{C}^n || \underline{x} - \underline{\tilde{s}}| < \rho\}, \quad \rho > 0$ such that

$$\inf_{\zeta \in K} |\phi(\zeta, \underline{a})| > 0, \quad \forall \underline{a} \in B.$$

By the principle of argument, the number of zeros of $\phi(z,\underline{a})$ within K is given by

$$N(\underline{a}) = \frac{1}{2\pi i}\oint_K \frac{\phi'(z,\underline{a})}{\phi(z,\underline{a})}dz, \quad \phi' = \frac{\partial \phi}{\partial z}$$

which is continuous in B; hence

$$1 = N(\underline{\tilde{s}}) = N(\underline{a}), \quad \underline{a} \in B.$$

Moreover the simple zero $\xi(\emptyset)$ of $\phi(z,\underline{a})$ inside K admits the representation (see e.g. [21])

$$\xi(\emptyset) = \frac{1}{2\pi i}\oint_K \frac{z\phi'(z,\underline{a})}{\phi(z,\underline{a})}dz.$$

For $\underline{a} \in B$ the integrand is an analytic function of $\underline{a}$ and therefore also $\xi(\emptyset)$ is an analytic function of $\underline{a}$ when $\underline{a} \in B$.

We now consider $T^{(2)}(\underline{a})$. We notice that each component can be obtained as a rational function of the components of $T^{(1)}(\underline{a})$ by the formula $c_j = \underline{e}_j^T V^{-H}\underline{a}$, $j = 1,\ldots,n/2$ where V is the Vandermonde matrix based on $T^{(1)}(\underline{a})$. Therefore also c_j is an analytic function of $\underline{a}$ when $\underline{a} \in B$.

As $T^{(h)} = T_R^{(h)} + iT_I^{(h)}$ is analytic for $\underline{a} \in B$, $T_R^{(h)}$ and $T_I^{(h)}$ are real analytic functions of $\underline{a}_R, \underline{a}_I$ where $\underline{a} = \underline{a}_R + i\underline{a}_I$, (e.g. [13][pg.99]). Therefore they admit a Taylor series expansion around $\underline{\tilde{s}}$ when $\underline{a} \in B$:

$$\begin{aligned} T_{Rk}^{(h)}(\underline{a}) = T_{Rk}^{(h)}(\underline{\tilde{s}}) + \sum_{i=0}^{n-1} \frac{\partial T_{Rk}^{(h)}(\underline{a})}{\partial a_{Ri}}_{|\underline{a}=\underline{\tilde{s}}} [a_{Ri} - \tilde{s}_{Ri}] + \\ \sum_{i=0}^{n-1} \frac{\partial T_{Rk}^{(h)}(\underline{a})}{\partial a_{Ii}}_{|\underline{a}=\underline{\tilde{s}}} [a_{Ii} - \tilde{s}_{Ii}] + \\ \frac{1}{2}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} \frac{\partial^2 T_{Rk}^{(h)}(\underline{a})}{\partial a_{Ri}\partial a_{Rj}}_{|\underline{a}=\underline{\tilde{s}}} [a_{Ri} - \tilde{s}_{Ri}][a_{Rj} - \tilde{s}_{Rj}] + \end{aligned}$$

$$\frac{1}{2}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\frac{\partial^2 T_{Rk}^{(h)}(\underline{a})}{\partial a_{Ii}\partial a_{Ij}}_{|\underline{a}=\underline{\tilde{s}}}[a_{Ii}-\tilde{s}_{Ii}][a_{Ij}-\tilde{s}_{Ij}]+$$
$$\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\frac{\partial^2 T_{Rk}^{(h)}(\underline{a})}{\partial a_{Ri}\partial a_{Ij}}_{|\underline{a}=\underline{\tilde{s}}}[a_{Ri}-\tilde{s}_{Ri}][a_{Ij}-\tilde{s}_{Ij}]+...$$

and analogously for $T_{Ik}^{(h)}(\underline{a})$. Taking expectations we get

$$E\left[(\mathbf{a}_{Ri}-\tilde{s}_{Ri})\right]=[s_{Ri}-\tilde{s}_{Ri}]=\sigma^\alpha\cdot C_i,\quad C_i=\sum_{j=p+1}^{n/2}\tilde{\xi}_{j-p}^i$$

$$E\left[(\mathbf{a}_{Ri}-\tilde{s}_{Ri})(\mathbf{a}_{Rj}-\tilde{s}_{Rj})\right]=E\left[(\mathbf{a}_{Ri}-s_{Ri}+\sigma^\alpha C_i)(\mathbf{a}_{Rj}-s_{Rj}+\sigma^\alpha C_j)\right]$$
$$=\frac{\sigma^2}{2}\delta_{ij}+\sigma^{2\alpha}C_iC_j$$

and analogously for the other terms. Remembering the independence of the real and imaginary parts of $\mathbf{a}_k$, we finally get

$$E[T_k^{(h)}(\underline{\mathbf{a}})]=T_k^{(h)}(\underline{\tilde{s}})+o(\sigma).\qquad\square$$

We start now the study of the distribution in $\mathbb{C}$ of the generalized eigenvalues of $\mathbf{P}$ by making some qualitative statements already present in the literature. For each realization ø, let $\{c_j(ø),\xi_j(ø)\},\ j=1,\ldots,n/2$ be the solution of the complex exponential interpolation problem for the data $a_k(ø),\ k=0,\ldots,n-1$. It is well known that we can then define the Pade' approximant

$$[n/2-1,n/2](z,ø)=z\sum_{j=1}^{n/2}\frac{c_j(ø)}{z-\xi_j(ø)}=Q_{n/2-1}(z^{-1})/P_{n/2}(z^{-1})$$

to the $Z-$transform of $\{a_k(ø)\}$ given by

$$f(z,ø)=\sum_{k=0}^{\infty}a_k(ø)z^{-k}=f_s(z)+f_\nu(z,ø)$$

where

$$f_s(z) = \sum_{k=0}^{\infty} s_k z^{-k} = \sum_{j=1}^{p} c_j \sum_{k=0}^{\infty} (\xi_j/z)^k = z \sum_{j=1}^{p} \frac{c_j}{z-\xi_j}, \quad |z| > 1$$

and, because of Lemma 1,

$$f_\nu(z, \o) \approx z \sum_{j=1}^{n_p} \frac{\tilde{c}_j(\o)}{z - \tilde{\xi}_j}$$

$f(z, \o)$ is then defined outside the unit circle and can be extended to D by analytic continuation. We get then

$$f(z, \o) \approx z\tilde{q}_{n/2-1}(z)/\tilde{p}_{n/2}(z) = \frac{z \prod_{j=1}^{n/2-1}(z - \delta_j(\o))}{\prod_{j=1}^{p}(z - \xi_j) \prod_{j=1}^{n_p}(z - \tilde{\xi}_j)}$$

and

$$g(z, \o) = \log(z^{-1} f(z, \o)) = \sum_{j=1}^{n/2-1} \log(z - \delta_j(\o)) - \sum_{j=1}^{p} \log(z - \xi_j) - \sum_{j=1}^{n_p} \log(z - \tilde{\xi}_j).$$

We want to study the location in $\mathbb{C}$ of $\xi_j(\o)$. To this aim, following [19], we remember that $p_n(z) = z^n P_n(z^{-1})$ satisfy the following orthogonality relation

$$\int_\Gamma z^{-1} f(z, \o) p_n(z) z^k dz = 0, \quad k = 0, \ldots, n-1$$

where Γ is a union of closed curves enclosing the poles of $f(z, \o)$ i.e. the numbers $\xi_j, \; j = 1 \ldots, p$ and $\tilde{\xi}_j, \quad j = 1, \ldots, n_p$. By using the Szego integral representation of such polynomials and a saddle point argument, it turns out that the Pade' poles $\xi_j(\o), \; j = 1, \ldots, n/2$, asymptotically on n, satisfy the following system of algebraic equations

$$2 \sum_{j \neq k}^{1, n/2} \frac{1}{(\xi_k(\o) - \xi_j(\o))} + g'(\xi_k(\o)) = 0 \qquad\qquad k = 1, \ldots, n/2$$

or

$$2\sum_{j\neq k}^{1,n/2}\frac{1}{(\xi_k(\o)-\xi_j(\o))}+\sum_{j=1}^{n/2-1}\frac{1}{(\xi_k(\o)-\delta_j(\o))}+$$
$$-\sum_{j=1}^{p}\frac{1}{(\xi_k(\o)-\xi_j)}-\sum_{j=1}^{n_p}\frac{1}{(\xi_k(\o)-\tilde{\xi}_j)}=0,\quad k=1,\ldots,n/2$$

These equations can be interpreted as conditions of electrostatic equilibrium of a set of charges in the presence of an electric external field corresponding to $g'(z,\o)$. Therefore the Pade' poles $\xi_k(\o)$ are attracted by $\xi_j,\ j=1,\ldots,p$ and $\tilde{\xi}_j,\ j=1,\ldots n_p$ and they are repelled by each other and by the zeros $\delta_j(\o)$ of $\tilde{q}_{n/2-1}(z)$. However

$$\tilde{q}_{n/2-1}(z)=\sum_{j=1}^{p}c_j\prod_{k\neq j}^{1,p}(z-\xi_k)\prod_{k=1}^{n_p}(z-\tilde{\xi}_k) \tag{3}$$
$$+\sum_{j=1}^{n_p}\tilde{c}_j(\o)\prod_{k=1}^{p}(z-\xi_k)\prod_{k\neq j}^{1,n_p}(z-\tilde{\xi}_k). \tag{4}$$

As $\forall\o,\ \ |\tilde{c}_j(\o)|^2\ll\min_h|c_h|^2$ if the SNR is sufficiently high (see Remark after Lemma 1), we can approximate $\tilde{q}_{n/2-1}(z)$ by

$$\prod_{k=1}^{n_p}(z-\tilde{\xi}_k)\sum_{j=1}^{p}c_j\prod_{k\neq j}^{1,p}(z-\xi_k)$$

hence n_p zeros are close to $\tilde{\xi}_k$, and the other $p-1$ are close to the zeros of the polynomial

$$q_{p-1}(z)=\sum_{j=1}^{p}c_j\prod_{k\neq j}^{1,p}(z-\xi_k)$$

which is the numerator of $z^{-1}f_s(z)$. We notice that if $|c_h|\ll|c_k|,\ \forall k\neq h$ then

$$q_{p-1}(z)\approx\sum_{j\neq h}^{1,p}c_j\prod_{k\neq j}^{1,p}(z-\xi_k)=(z-\xi_h)\sum_{j=1}^{p}c_j\prod_{k\neq j,h}^{1,p}(z-\xi_k)$$

Hence, because of the continuous dependence of the roots from the coefficient of a polynomial, $q_{p-1}(z)$ has a zero as close to ξ_h as $|c_h|$ is small with respect to $|c_k|,\ k\neq h$. Therefore the Pade' poles $\xi_k(\o)$

- are attracted by $\xi_j,\ j = 1, \ldots, p$
- are attracted by $\tilde{\xi}_j,\ j = 1, \ldots n_p$
- are repelled from $\xi_j(\o),\ j \neq k$
- are repelled from $\tilde{\xi}_j,\ j = 1, \ldots n_p$
- are repelled from other $p-1$ points in the complex plane which are as close to ξ_j as $|c_j|$ is small with respect to $|c_h|,\ h \neq j$.

Summing up a ξ_k with a large $|c_k|$ will attract a Pade' pole without being disturbed by the repulsion exerted by the zeros of $\tilde{q}_{n/2-1}(z)$. Moreover close to such a point a gap of Pade' poles can be expected because of the repulsion exerted by Pade' poles to each other. A ξ_k with a small $|c_k|$ will still attract a Pade' pole but not so close because of the repulsion exerted by a close zero. The Pade' poles not related to the signal are expected to be attracted by $\tilde{\xi}_k$ which at the same time will repel them. Moreover they are repelled by ξ_k hence they are likely to be located in between $\tilde{\xi}_k$ and far from ξ_k. A picture of this behavior is given in fig.1. We notice that the qualitative results discussed above are consistent with those obtained in [3] under a more stringent hypothesis about the noise.

We now wish to define a mathematical tool to quantify these qualitative statements. To this aim we remember that $\boldsymbol{\xi}_k, k = 1, \ldots, n/2$ are the generalized eigenvalues of the pencil $\mathbf{P}$ and therefore they satisfy the equation

$$\mathbf{P}_{n/2}(z^{-1}) = det[U_1(\underline{\mathbf{a}}) - zU_0(\underline{\mathbf{a}})] = 0.$$

Then a condensed density $h_n(z)$ can be considered which is the expected value of the (random) normalized counting measure on the zeros of this polynomial i.e.

$$h_n(z) = \frac{2}{n} E\left[\sum_{j=1}^{n/2} \delta(z - \boldsymbol{\xi}_j)\right].$$

The following theorem holds whose proof is the same of that of Theorem 1 in [1]:

Theorem 1 *The condensed density of the zeros of the random polynomial* $\mathbf{Q}(z) = \mathbf{P}_{n/2}(z^{-1})$ *is given by*

$$h_n(z) = \frac{1}{4\pi} \Delta u_n(z) \tag{5}$$

where Δ *denotes the Laplacian operator with respect to* x, y *if* $z = x + iy$ *and*

$$u_n(z) = \frac{2}{n} E\left\{\log(|\mathbf{Q}(z)|^2)\right\} \tag{6}$$

The condensed density provides the required quantitative information about the distribution of the Pade' poles in the complex plane. If the SNR is sufficiently high, after the qualitative statements made above about the location of the Pade' poles, a peak of $h_n(z)$ can be expected in a neighborhood of each of the complex exponentials $\xi_k, k = 1, \ldots, p$ and the volume under the peak gives the probability of finding a Pade' pole in that neighborhood. This is confirmed by the following

Theorem 2 *If* $\sigma > 0$*, the condensed density* $h_n(z, \sigma)$ *is a continuous function of* z *given by*

$$h_n(z,\sigma) = \frac{2}{n(\pi\sigma^2)^n} \sum_{j=1}^{n/2} \int_{\mathcal{C}^{n/2-1}} \int_{\mathcal{C}^{n/2}} J_C^*(\underline{\zeta}_j^*, z, \underline{\gamma}) e^{-\frac{1}{\sigma^2}\sum_{k=0}^{n-1} |\sum_{h\neq j}^{1,n/2} \gamma_h \zeta_h^k + \gamma_j z^k - s_k|^2} d\underline{\zeta}_j^* d\underline{\gamma} \tag{7}$$

where $\underline{\zeta}_j^* = \{\zeta_h, h \neq j\}$ *and*

$$J_C^*(\underline{\zeta}_j^*, z, \underline{\gamma}) = \begin{cases} \gamma & \text{if } n = 2 \\ (-1)^{n/2} \prod_{j=1}^{1,n/2} \gamma_j \prod_{r<h, r\neq j} (\zeta_r - \zeta_h)^4 \prod_{r \neq j} (\zeta_r - z)^4 & \text{if } n \geq 4 \end{cases}$$

Moreover $h_n(z, \sigma)$ *converges weakly to the positive measure* $\frac{2}{n} \sum_{j=1}^{p} \delta(z - \xi_j)$ *when* $\sigma \to 0$.

proof

Let us consider the transformation $T_n : \underline{\alpha} \to (\underline{\zeta}, \underline{\gamma})$ given by

$$\alpha_k = \sum_{j=1}^{n/2} \gamma_j \zeta_j^k$$

or

$$(T_n^{(1)}(\underline{\alpha}))_j = \zeta_j, \quad (T_n^{(2)}(\underline{\alpha}))_j = \gamma_j.$$

In the following, to simplify notations, $(T_n^{(1)}(\underline{\alpha}))_j$ will be denoted by $\zeta_j(\underline{\alpha})$. We have

$$h_n(z, \sigma) = \frac{2}{n} E\left[\sum_{j=1}^{n/2} \delta(z - \boldsymbol{\xi}_j)\right] \tag{8}$$

$$= \frac{2}{n(\pi\sigma^2)^n} \sum_{j=1}^{n/2} \int_{\mathbb{C}^n} \delta(z - \zeta_j(\underline{\alpha}) e^{-\frac{1}{\sigma^2} \sum_{k=0}^{n-1} |\alpha_k - s_k|^2} d\underline{\alpha}; \tag{9}$$

As the complex Jacobian of T_n^{-1} is (see [9,17]) (n was assumed even):

$$J_C(\underline{\zeta}, \underline{\gamma}) = \begin{cases} \gamma & \text{if } n = 2 \\ (-1)^{n/2} \prod_{j=1}^{n/2} \gamma_j \prod_{j<h} (\zeta_j - \zeta_h)^4 & \text{if } n \geq 4 \end{cases},$$

by making a change of variables we have

$$h_n(z,\sigma)=\frac{2}{n(\pi\sigma^2)^n}\sum_{j=1}^{n/2}\int_{\mathbb{C}^{n/2}}\int_{\mathbb{C}^{n/2}}\delta(z-\zeta_j)J_C(\underline{\zeta},\underline{\gamma})e^{-\frac{1}{\sigma^2}\sum_{k=0}^{n-1}|\sum_{h=1}^{n/2}\gamma_h\zeta_h^k-s_k|^2}d\underline{\zeta}d\underline{\gamma}$$

$$=\frac{2}{n(\pi\sigma^2)^n}\sum_{j=1}^{n/2}\int_{\mathbb{C}^{n/2-1}}\int_{\mathbb{C}^{n/2}}J_C^*(\underline{\zeta}_j^*,z,\underline{\gamma})e^{-\frac{1}{\sigma^2}\sum_{k=0}^{n-1}|\sum_{h\neq j}^{1,n/2}\gamma_h\zeta_h^k+\gamma_j z^k-s_k|^2}d\underline{\zeta}_j^*d\underline{\gamma}$$

where $\underline{\zeta}_j^*=\{\zeta_h,h\neq j\}$ and

$$J_C^*(\underline{\zeta}_j^*,z,\underline{\gamma})=\begin{cases}\gamma & \text{if } n=2\\ (-1)^{n/2}\prod_{j=1}^{n/2}\gamma_j\prod_{r<h,r\neq j}(\zeta_r-\zeta_h)^4\prod_{r\neq j}(\zeta_r-z)^4 & \text{if } n\geq 4\end{cases}$$

The integral above converges uniformly for $z\in D$, hence $h_n(z)$ is continuous in D. We prove now that $h_{2p}(z,\sigma)$ converges weakly to $\frac{1}{p}\sum_{j=1}^{p}\delta(z-\xi_j)$ when $\sigma\to 0$. Let $\Phi(z)\in C^\infty$ be a bounded test function supported on $\mathbb{C}$. We have

$$\int_{\mathbb{C}} h_{2p}(z,\sigma)\Phi(z)dz$$
$$=\frac{1}{p(\pi\sigma^2)^{2p}}\sum_{j=1}^{p}\int_{\mathbb{C}}\Phi(z)\left[\int_{\mathbb{C}^{2p}}\delta(z-\zeta_j(\underline{\alpha}))e^{-\frac{1}{\sigma^2}\sum_{k=0}^{2p-1}|\alpha_k-s_k|^2}d\underline{\alpha}\right]dz$$
$$=\frac{1}{p(\pi\sigma^2)^{2p}}\sum_{j=1}^{p}\int_{\mathbb{C}^{2p}}\Phi(\zeta_j(\underline{\alpha}))e^{-\frac{1}{\sigma^2}\sum_{k=0}^{2p-1}|\alpha_k-s_k|^2}d\underline{\alpha}$$
$$=\frac{1}{p}\sum_{j=1}^{p}\int_{\mathbb{C}^{2p}}\Phi(\zeta_j(\underline{y}\sigma+\underline{s}))\frac{e^{-\sum_{k=0}^{2p-1}|\underline{y}_k|^2}}{\pi^{2p}}d\underline{y}.$$

As $\Phi(z)$ is continuous and bounded and ζ_j is analytic in a neighbor of $\underline{s}$ by Lemma 2, by the dominated convergence theorem we get

$$\lim_{\sigma\to 0}\int_{\Omega} h_{2p}(z,\sigma)\Phi(z)dz=\frac{1}{p}\sum_{j=1}^{p}\int_{\mathbb{C}^{2p}}\lim_{\sigma\to 0}\Phi(\zeta_j(\underline{y}\sigma+\underline{s}))\frac{e^{-\sum_{k=0}^{2p-1}|\underline{y}_k|^2}}{\pi^{2p}}d\underline{y}=$$
$$\frac{1}{p}\sum_{j=1}^{p}\Phi(\zeta_j(\underline{s}))\int_{\mathbb{C}^{2p}}\frac{e^{-\sum_{k=0}^{2p-1}|\underline{y}_k|^2}}{\pi^{2p}}d\underline{y}=\frac{1}{p}\sum_{j=1}^{p}\Phi(\zeta_j(\underline{s}))=\frac{1}{p}\sum_{j=1}^{p}\Phi(\xi_j)$$

because $(T_{2p}^{(1)}(\underline{s}))_j = \xi_j$.

Let us consider now the case $n > 2p$. We cannot use the same argument used for the case $n = 2p$ because $\zeta_j(\underline{s})$ is not defined for $j = p+1, \ldots, n/2$ (see Lemma 2). However by Lemma 1 without loss of generality, we can assume that the noise is represented by $\tilde{\boldsymbol{\nu}}_k$ i.e.

$$\mathbf{a}_k = \begin{cases} \sum_{j=1}^{p} c_j \xi_j^k + \sum_{j=p+1}^{n/2} \tilde{\mathbf{c}}_{j-p} \tilde{\xi}_{j-p}^k, & k = 0, \ldots, n_p - 1 \\ \\ \sum_{j=1}^{p} c_j \xi_j^k + \boldsymbol{\nu}_k, & k = n_p, \ldots, n-1 \end{cases}$$

where $n_p = n/2 - p$. We then define a new process $\tilde{\mathbf{a}}_k$ by

$$\tilde{\mathbf{a}}_k = \sum_{j=1}^{p} c_j \xi_j^k + \boldsymbol{\eta}_k, \quad k = 0, \ldots, n-1$$

where

$$\boldsymbol{\eta}_k = \sum_{j=p+1}^{n/2} \tilde{\mathbf{c}}_{j-p} \tilde{\xi}_{j-p}^k,$$

and we consider the process $\mathbf{a}_k$ as a perturbation of the process $\tilde{\mathbf{a}}_k$. Let us consider the pencils

$$\mathbf{P} = [U(\mathbf{a}_1, \ldots, \mathbf{a}_{n-1}), U(\mathbf{a}_0, \ldots, \mathbf{a}_{n-2})]$$

and

$$\tilde{\mathbf{P}} = [U(\tilde{\mathbf{a}}_1, \ldots, \tilde{\mathbf{a}}_{n-1}), U(\tilde{\mathbf{a}}_0, \ldots, \tilde{\mathbf{a}}_{n-2})].$$

We can write

$$\mathbf{P} = \tilde{\mathbf{P}} + \sigma \mathbf{E}$$

where

$$\mathbf{E} = \frac{1}{\sigma}[U(0, \ldots, 0, \boldsymbol{\nu}_{n_p+1} - \boldsymbol{\eta}_{n_p+1}, \ldots, \boldsymbol{\nu}_{n-1} - \boldsymbol{\eta}_{n-1}),$$
$$U(0, \ldots, 0, \boldsymbol{\nu}_{n_p} - \boldsymbol{\eta}_{n_p}, \ldots, \boldsymbol{\nu}_{n-2} - \boldsymbol{\eta}_{n-2})] = [\mathbf{E}_1, \mathbf{E}_0].$$

From [16], in the limit for $\sigma \to 0$, a generalized eigenvalue $\boldsymbol{\xi}_j$ of $\mathbf{P}$ can be expressed as a function of a generalized eigenvalue $\hat{\xi}_j$ of $\tilde{\mathbf{P}}$ and corresponding left and right generalized eigenvectors $\underline{v}_j, \underline{u}_j$ by

$$\begin{aligned}\boldsymbol{\xi}_j &= \hat{\xi}_j + \sigma \frac{\underline{v}_j^H(\mathbf{E}_1 - \hat{\xi}_j \mathbf{E}_0)\underline{u}_j}{\underline{v}_j^H \mathbf{U}_0 \underline{u}_j} + O(\sigma^2) \\ &= \hat{\xi}_j + \sigma \frac{\underline{e}_j^T V^{-1}(\mathbf{E}_1 - \hat{\xi}_j \mathbf{E}_0))V^{-T}\underline{e}_j}{\hat{\mathbf{c}}_j} + O(\sigma^2)\end{aligned}$$

where $\mathbf{U}_0 = U(\tilde{\mathbf{a}}_1, \ldots, \tilde{\mathbf{a}}_{n-1})$ and, by construction,

$$\hat{\xi}_j = \begin{cases} \xi_j & j = 1, \ldots, p \\ \\ \tilde{\xi}_{j-p}, & j = p+1, \ldots, n/2 \end{cases}$$

$$\hat{\mathbf{c}}_j = \begin{cases} c_j & j = 1, \ldots, p \\ \\ \tilde{\mathbf{c}}_{j-p}, & j = p+1, \ldots, n/2 \end{cases}$$

$$V = Vander(\hat{\xi}_1, \ldots, \hat{\xi}_{n/2}), \quad C = diag(\hat{\mathbf{c}}_1, \ldots, \hat{\mathbf{c}}_{n/2})$$

and

$$\underline{v}_j = \overline{\underline{u}}_j = V^{-H}\underline{e}_j.$$

We notice that we can write

$$\underline{e}_j^T V^{-1}(\mathbf{E}_1 - \hat{\xi}_j \mathbf{E}_0)V^{-T}\underline{e}_j = \sum_{h=1}^{n/2+p} \gamma_{jh} \mathbf{Y}_{n_p+h}$$

where γ_{jh} are constants and $\mathbf{Y}_h($ are i.i.d. zero mean, complex Gaussian variables with unit variance identified with $\frac{1}{\sqrt{2}\sigma}[\boldsymbol{\nu}_h - \boldsymbol{\eta}_h],\ h = n_p, \ldots, n-1$.

We have

$$
\begin{aligned}
h_n(z,\sigma) &= \frac{2}{n}E\left[\sum_{j=1}^{n/2}\delta(z-\boldsymbol{\xi}_j)\right] = \frac{2}{n}E\left[\sum_{j=1}^{p}\delta(z-\boldsymbol{\xi}_j)\right] + \frac{2}{n}E\left[\sum_{j=p+1}^{n/2}\delta(z-\boldsymbol{\xi}_j)\right] \\
&= h_n^{(1)}(z,\sigma) + h_n^{(2)}(z,\sigma)
\end{aligned}
$$

By the same argument used for the case $n = 2p$ it follows that $h_n^{(1)}(z,\sigma)$ converges weakly to $\frac{2}{n}\sum_{j=1}^{p}\delta(z-\xi_j)$ when $\sigma \to 0$. We then consider $h_n^{(2)}(z,\sigma)$. We have

$$
\begin{aligned}
h_n^{(2)}(z,\sigma) &= \frac{2}{n}E\left[\sum_{j=p+1}^{n/2}\delta(z-\boldsymbol{\xi}_j)\right] \\
&= \frac{2}{n}E\left[\sum_{j=p+1}^{n/2}\delta\left(z-\tilde{\xi}_{j-p}-\sigma\frac{\sum_{h=1}^{n/2+p}\gamma_{jh}\mathbf{Y}_{n_p+h}}{\tilde{\mathbf{c}}_{j-p}}-O(\sigma^2)\right)\right].
\end{aligned}
$$

By identifying $\frac{\sqrt{n_p}}{\sigma}\tilde{\mathbf{c}}_{j-p}, \;\; j = p+1,\ldots,n/2$ with $\mathbf{Y}_h, \;\; h = 1,\ldots,n_p$, which are i.i.d. zero mean, complex Gaussian variables with unit variance, we get

$$
\begin{aligned}
h_n^{(2)}(z,\sigma) &= \sum_{j=p+1}^{n/2}\int_{\mathbb{C}^n}\delta\left(z-\tilde{\xi}_{j-p}-\frac{\sqrt{n_p}}{y_{j-p}}\sum_{h=1}^{n/2+p}\gamma_{jh}y_{n_p+h}-O(\sigma^2)\right)\frac{e^{-\frac{1}{\sigma^2}\sum_{k=1}^{n}|y_k|^2}}{\pi^n}d\underline{y} \\
&= \sum_{j=p+1}^{n/2}\int_{\mathbb{C}^{n-1}}\left[\int_{\mathbb{C}}\delta\left(z-\tilde{\xi}_{j-p}-\frac{\sqrt{n_p}}{y_{j-p}}\sum_{h=1}^{n/2+p}\gamma_{jh}y_{n_p+h}-O(\sigma^2)\right)\frac{e^{-|y_{j-p}|^2}}{\pi}dy_{j-p}\right] \\
&\quad\cdot\frac{e^{-\sum_{k=1,k\neq j-p}^{n}|y_k|^2}}{\pi^{n-1}}d\underline{y}', \quad \{\underline{y}'\} = \{\underline{y}\}\setminus\{y_{j-p}\}
\end{aligned}
\tag{10}
$$

by making the change of variable

$$
w = \tilde{\xi}_{j-p} + \frac{\sqrt{n_p}}{y_{j-p}}\sum_{h=1}^{n/2+p}\gamma_{jh}y_{n_p+h}
$$

we get

$$
\int_{\mathbb{C}}\delta\left(z-\tilde{\xi}_{j-p}-\frac{\sqrt{n_p}}{y_{j-p}}\sum_{h=1}^{n/2+p}\gamma_{jh}y_{n_p+h}-O(\sigma^2)\right)\frac{e^{-|y_{j-p}|^2}}{\pi}dy_{j-p}
$$

$$
= -\frac{1}{\pi}\int_{\mathbb{C}} \delta\left(z - w - O(\sigma^2)\right) \frac{\sqrt{n_p}\sum_{h=1}^{n/2+p}\gamma_{jh}y_{n_p+h}}{(w-\tilde{\xi}_{j-p})^2} e^{-\left|\frac{\sqrt{n_p}\sum_{h=1}^{n/2+p}\gamma_{jh}y_{n_p+h}}{w-\tilde{\xi}_{j-p}}\right|^2} dw
$$

$$
= -\frac{1}{\pi}\frac{\sqrt{n_p}\sum_{h=1}^{n/2+p}\gamma_{jh}y_{n_p+h}}{(z-O(\sigma^2)-\tilde{\xi}_{j-p})^2} e^{-\left|\frac{\sqrt{n_p}\sum_{h=1}^{n/2+p}\gamma_{jh}y_{n_p+h}}{z-O(\sigma^2)-\tilde{\xi}_{j-p}}\right|^2}.
$$

Inserting this expression in (10) we get

$$
h_n^{(2)}(z,\sigma) = -\sum_{j=p+1}^{n/2} \frac{\sqrt{n_p}}{(z-O(\sigma^2)-\tilde{\xi}_{j-p})^2} \cdot \sum_{r=1}^{n/2+p}\gamma_{jr}\frac{1}{\pi^n}\int_{\mathbb{C}^{n-1}} y_{n_p+r} e^{-\left|\frac{\sqrt{n_p}\sum_{h=1}^{n/2+p}\gamma_{jh}y_{n_p+h}}{z-O(\sigma^2)-\tilde{\xi}_{j-p}}\right|^2 - \sum_{k=1,k\neq j-p}^{n}|y_k|^2} d\underline{y}'
$$

and therefore

$$
\lim_{\sigma\to 0} h_n^{(2)}(z,\sigma) = -\sum_{j=p+1}^{n/2} \frac{\sqrt{n_p}}{(z-\tilde{\xi}_{j-p})^2} \cdot \sum_{r=1}^{n/2+p}\gamma_{jr}\frac{1}{\pi^n}\int_{\mathbb{C}^{n-1}} y_{n_p+r} e^{-\left|\frac{\sqrt{n_p}\sum_{h=1}^{n/2+p}\gamma_{jh}y_{n_p+h}}{z-\tilde{\xi}_{j-p}}\right|^2 - \sum_{k=1,k\neq j-p}^{n}|y_k|^2} d\underline{y}' = 0
$$

because

$$
\frac{1}{\pi^{n-1}}\int_{\mathbb{C}^{n-1}} y_{n_p+r} e^{-\left|\frac{\sqrt{n_p}\sum_{h=1}^{n/2+p}\gamma_{jh}y_{n_p+h}}{z-\tilde{\xi}_{j-p}}\right|^2 - \sum_{k=1,k\neq j-p}^{n}|y_k|^2} d\underline{y}'
$$

$$
= \frac{1}{\pi^{n-1}}\int_{\mathbb{C}^{n-1}} y_{n_p+r} e^{-\underline{y}'^H A \underline{y}'} d\underline{y}' = 0, \quad \text{for a suitable hermitian matrix } A, \ \forall r. \qquad \square
$$

Remark. When the SNR is large the exponential part dominates the integrand as the Jacobian does not depend on σ. Moreover the exponential part has relative maxima close to ξ_j as expected. In general the integral (7) does not admit a closed form expression. However when $n = 2$, remembering that the Jacobian with respect

to the real and imaginary part of a complex variable is $J_R = |J_C|^2$, the integral (7) becomes

$$h_2(z,\sigma) = \frac{1}{(\pi\sigma^2)^2} \int_{\mathbb{C}} \gamma e^{-\frac{|\gamma - s_0|^2 + |\gamma z - s_1|^2}{\sigma^2}} d\gamma = \frac{1}{(\pi\sigma^2)^2} \int_{\mathbb{R}^2} |\gamma|^2 e^{-\frac{|\gamma - s_0|^2 + |\gamma z - s_1|^2}{\sigma^2}} d\Re\gamma d\Im\gamma$$
$$= \frac{\sigma^2(1+|z|^2) + |zs_1 + s_0|^2}{\pi\sigma^2(1+|z|^2)^3} e^{-\frac{|zs_0 - s_1|^2}{\sigma^2(1+|z|^2)}}.$$

We notice that $\lim_{\sigma\to 0} h_2(z,\sigma) = \delta(z - s_1/s_0) = \delta(z - \xi_1)$. Moreover, when $s_0 = s_1 = 0$ we have $h_2(z,\sigma) = \frac{1}{\pi(1+|z|^2)^2}$ which is independent of σ^2, confirming the result obtained in [1] for the pure noise case.

The condensed density has an important role in the following. Therefore we look for an easily computable approximation. The following theorem provides a basis for building such an approximation :

Theorem 3 *Let be* $\mathbf{F}(z,\overline{z}) = (U_1(\underline{\mathbf{a}}) - zU_0(\underline{\mathbf{a}}))\overline{(U_1(\underline{\mathbf{a}}) - zU_0(\underline{\mathbf{a}}))}$ *then*

$$E[\log(det\{\mathbf{F}(z,\overline{z})\})] - \log(det\{E[\mathbf{F}(z,\overline{z})]\}) = o(\sigma)$$

for $\sigma \to 0$*, independently of* z*. Moreover*

$$E[\mathbf{F}(z,\overline{z})] = (U_1(\underline{s}) - zU_0(\underline{s}))\overline{(U_1(\underline{s}) - zU_0(\underline{s}))} + \frac{n\sigma^2}{2} A(z,\overline{z}) \tag{11}$$

where

$$A(z,\overline{z}) = \begin{bmatrix} 1+|z|^2 & -z & 0 & \dots & 0 \\ -\overline{z} & 1+|z|^2 & -z & 0 & \dots \\ . & . & . & . & . \\ 0 & \dots & 0 & -\overline{z} & 1+|z|^2 \end{bmatrix}.$$

<u>proof</u>

let us denote by $\boldsymbol{\lambda}_j$ the eigenvalues of $\mathbf{F}(z,\overline{z})$ and by μ_j those of $E[\mathbf{F}(z,\overline{z})]$, dropping for simplicity the dependence on $z,\overline{z}$. Note that $\mu_j \neq E[\boldsymbol{\lambda}_j]$, see e.g. [5, Theorem 8.5]. We have

$$E[\log(det\{\mathbf{F}(z,\overline{z})\})] = \sum_j E[\log(\boldsymbol{\lambda}_j)]$$

and

$$\log(det\{E[\mathbf{F}(z,\overline{z})]\}) = \sum_j \log(\mu_j)$$

hence it is sufficient to study the difference

$$E[\log(\boldsymbol{\lambda}_j)] - \log(\mu_j).$$

We then denote by $\underline{\mathbf{f}}$ the vector obtained by stacking the real and imaginary parts of the elements $(\mathbf{F}_{hk}, h,k = 1,\dots,n/2)$ of $\mathbf{F}$ and consider the function

$$g(\underline{\mathbf{f}}) = \log(\boldsymbol{\lambda}_j)$$

and its Taylor expansion around $E[\underline{\mathbf{f}}]$:

$$
\begin{aligned}
g(\underline{\mathbf{f}}) = {} & g(E[\underline{\mathbf{f}}]) + \sum_h \frac{\partial g}{\partial \underline{\mathbf{f}}_h} \Big|_{E[\underline{\mathbf{f}}]} (\underline{\mathbf{f}}_h - E[\underline{\mathbf{f}}_h]) \\
& + \frac{1}{2} \sum_{hk} \frac{\partial^2 g}{\partial \underline{\mathbf{f}}_h \partial \underline{\mathbf{f}}_k} \Big|_{E[\underline{\mathbf{f}}]} (\underline{\mathbf{f}}_h - E[\underline{\mathbf{f}}_h])(\underline{\mathbf{f}}_k - E[\underline{\mathbf{f}}_k]) + \dots
\end{aligned}
$$

which can be rewritten as

$$
\log(\boldsymbol{\lambda}_j) - \log(\mu_j) = \sum_h \beta_h (\underline{\mathbf{f}}_h - E[\underline{\mathbf{f}}_h]) + \frac{1}{2} \sum_{hk} \gamma_{hk} (\underline{\mathbf{f}}_h - E[\underline{\mathbf{f}}_h])(\underline{\mathbf{f}}_k - E[\underline{\mathbf{f}}_k]) + \dots
$$

and, taking expectations,

$$
E[\log(\boldsymbol{\lambda}_j)] - \log(\mu_j) = \frac{1}{2} \sum_{hk} \gamma_{hk} E[(\underline{\mathbf{f}}_h - E[\underline{\mathbf{f}}_h])(\underline{\mathbf{f}}_k - E[\underline{\mathbf{f}}_k])] + \dots
$$

But

$$
\begin{aligned}
\mathbf{F}(z, \overline{z}) = {} & (U_1(\underline{s}) - zU_0(\underline{s}) \overline{(U_1(\underline{s}) - zU_0(\underline{s}))} \\
& + (U_1(\underline{\boldsymbol{\nu}}) - zU_0(\underline{\boldsymbol{\nu}})) \overline{(U_1(\underline{\boldsymbol{\nu}}) - zU_0(\underline{\boldsymbol{\nu}}))} \\
& - (U_1(\underline{s}) - zU_0(\underline{s}) \overline{(U_1(\underline{\boldsymbol{\nu}}) - zU_0(\underline{\boldsymbol{\nu}}))} \\
& - (U_1(\underline{\boldsymbol{\nu}}) - zU_0(\underline{\boldsymbol{\nu}})) \overline{(U_1(\underline{s}) - zU_0(\underline{s}))}
\end{aligned}
$$

and

$$
\begin{aligned}
E[\mathbf{F}(z, \overline{z})] = {} & (U_1(\underline{s}) - zU_0(\underline{s}) \overline{(U_1(\underline{s}) - zU_0(\underline{s}))} \\
& + E[(U_1(\underline{\boldsymbol{\nu}}) - zU_0(\underline{\boldsymbol{\nu}})) \overline{(U_1(\underline{\boldsymbol{\nu}}) - zU_0(\underline{\boldsymbol{\nu}}))}] \\
= {} & (U_1(\underline{s}) - zU_0(\underline{s}) \overline{(U_1(\underline{s}) - zU_0(\underline{s}))} + \frac{n\sigma^2}{2} A(z, \overline{z})
\end{aligned}
$$

by a straightforward computation similar to that given in [1, Th.3] for the pure noise case. Therefore

$$
\begin{aligned}
\mathbf{F}(z, \overline{z}) - E[\mathbf{F}(z, \overline{z})] = {} & (U_1(\underline{\boldsymbol{\nu}}) - zU_0(\underline{\boldsymbol{\nu}})) \overline{(U_1(\underline{\boldsymbol{\nu}}) - zU_0(\underline{\boldsymbol{\nu}}))} \\
& - (U_1(\underline{s}) - zU_0(\underline{s}) \overline{(U_1(\underline{\boldsymbol{\nu}}) - zU_0(\underline{\boldsymbol{\nu}}))} \\
& - (U_1(\underline{\boldsymbol{\nu}}) - zU_0(\underline{\boldsymbol{\nu}})) \overline{(U_1(\underline{s}) - zU_0(\underline{s}))} \\
& - \frac{n\sigma^2}{2} A(z, \overline{z})
\end{aligned}
$$

hence $E[(\underline{\mathbf{f}}_h - E[\underline{\mathbf{f}}_h])(\underline{\mathbf{f}}_k - E[\underline{\mathbf{f}}_k])]$ is a linear combination of functions of z and $\overline{z}$ with coefficients equal to either σ^2 or σ^4 because the odd moments of a Gaussian are zero. By a similar argument all the dropped terms in the Taylor expansion above will depend on even powers of σ. Hence

$$E[\log(\boldsymbol{\lambda}_j)] - \log(\mu_j) = o(\sigma)$$

independently of $z, \overline{z}$. $\square$

By noticing that $|\mathbf{Q}(z)|^2 = det\{\mathbf{F}(z, \overline{z})\}$, an approximation of the condensed density is then given by

$$\tilde{h}_n(z, \sigma) = \frac{1}{2\pi n} \Delta \sum_{\mu_j(z)>0} \log(\mu_j(z))$$

where $\mu_j(z)$ are the eigenvalues of $E[\mathbf{F}(z, \overline{z})]$. Unfortunately $\tilde{h}_n(z, \sigma)$ is not a probability density as it can eventually assume negative values. However the following results hold

Theorem 4 *The function $\tilde{h}_n(z, \sigma)$ is continuous in σ and in z. In the limit cases $\sigma = 0$ and $\{c_k = 0, k = 1, \ldots, p\}$ it is given respectively by*

$$\tilde{h}_n(z, 0) = \frac{2}{n} \sum_{j=1}^{p} \delta(z - \xi_j)$$

and by

$$\tilde{h}_n(z, \sigma) = \frac{1}{4\pi} \Delta w_n(z)$$

where

$$w_n(z) = \frac{1}{n} \log \sum_{j=0}^{n} |z|^{2j}.$$

Moreover, in this second case, $\lim_{n\to\infty} \tilde{h}_n(z, \sigma) = \delta(|z| - 1)$.

<u>proof</u>

$\tilde{h}_n(z,\sigma)$ is continuous in σ and in z because of the continuous dependence of the eigenvalues on the elements of the corresponding matrix. When $\sigma = 0$, let $V \in \mathbb{C}^{n/2,p}$ be the Vandermonde matrix such that $U_0(\underline{s}) = VCV^T$ and $U_1(\underline{s}) = VCZV^T$. Let $V = QR$ be the QR decomposition of V. Then

$$E[\mathbf{F}(z,\overline{z})] = QRC(Z - zI)R^T Q^T \overline{QR}\overline{(Z - zI)}CR^H Q^H.$$

But $R = \begin{pmatrix} \tilde{R} \\ \\ 0 \end{pmatrix}$, therefore $R^T\overline{R} = \tilde{R}^T\overline{\tilde{R}}$; moreover $Q^T\overline{Q} = I$, hence the eigenvalues of $E[\mathbf{F}(z,\overline{z})]$ are the same of those of the matrix

$$RC(Z - zI)R^T\overline{R}\overline{(Z - zI)}CR^H = \begin{pmatrix} \tilde{R}C(Z - zI)\tilde{R}^T\overline{\tilde{R}}\overline{(Z - zI)}C\tilde{R}^H & 0 \\ \\ 0 & 0 \end{pmatrix}.$$

The non-zero eigenvalues of $E[\mathbf{F}(z,\overline{z})]$ are then the same of those of the matrix

$$\tilde{R}C(Z - zI)\tilde{R}^T\overline{\tilde{R}}\overline{(Z - zI)}C\tilde{R}^H.$$

We then have

$$\begin{aligned}
\tilde{h}_n(z,0) &= \frac{1}{2\pi n}\Delta \sum_{\mu_j(z)>0} \log(\mu_j(z)) \\
&= \frac{1}{2\pi n}\Delta \log\left(\prod_{j=1}^{p} |z - \xi_j|^2 \cdot |det(\tilde{R})|^4 \prod_{j=1}^{p} c_j^2\right) \\
&= \frac{2}{4\pi n}\sum_{j=1}^{p} \Delta \log|z - \xi_j|^2 = \frac{2}{n}\sum_{j=1}^{p} \delta(z - \xi_j)
\end{aligned}$$

because $\frac{1}{4\pi}\Delta\log(|z|^2) = \delta(z)$ (see e.g. [25, pg.47]). When $\{c_k = 0, k = 1, \ldots, p\}$

$$\tilde{h}_n(z,\sigma) = \frac{1}{2\pi n}\Delta \log(det\{A(z,\overline{z})\}) = \frac{1}{2\pi n}\Delta \log(\sum_{j=0}^{n} |z|^{2j}).$$

The last part of the thesis follows by the same argument used in the proof of Theorem 3 in [1]. $\square$

Corollary 2 $\tilde{h}_n(z,\sigma) - h_n(z,\sigma)$ *converges weakly to* 0 *when* $\sigma \to 0$

proof

Let $\Phi(z)$ be a nonnegative test function supported on $\mathbb{C}$. Denoting by $h_n^*(z) = \frac{2}{n}\sum_{j=1}^{p}\delta(z-\xi_j)$, from Theorems 2 and 4 we have $\forall \nu > 0$, $\exists \sigma_1$ and $\sigma_2 > 0$ such that

$$\left|\int_{\mathbb{C}} \Phi(z)\left(h_n(z,\sigma) - h_n^*(z)\right) dz\right| < \frac{\nu}{2}, \quad \forall \sigma < \sigma_1$$

and

$$\left|\int_{\mathbb{C}} \Phi(z)\left(\tilde{h}_n(z,\sigma) - h_n^*(z)\right) dz\right| < \frac{\nu}{2}, \quad \forall \sigma < \sigma_2$$

hence, if $\sigma_\nu = \min\{\sigma_1, \sigma_2\}$, we have $\forall \sigma < \sigma_\nu$

$$\left|\int_{\mathbb{C}} \Phi(z)\left(h_n(z,\sigma) - \tilde{h}_n(z,\sigma)\right) dz\right|$$
$$\leq \left|\int_{\mathbb{C}} \Phi(z)\left(h_n(z,\sigma) - h^*(z)\right) dz\right| + \left|\int_{\mathbb{C}} \Phi(z)\left(\tilde{h}_n(z,\sigma) - h^*(z)\right) dz\right| \leq \nu. \qquad \square$$

2 Identifiability of $S(z)$ and approximation properties of $E[\mathbf{S}_n(z)]$

We want now to exploit the information about the location in the complex plane of the Pade' poles, provided by the condensed density $h_n(z)$, to prove some properties relating $\mathbf{S}_n(z) = \sum_{j=1}^{n/2} \mathbf{c}_j \delta(z - \boldsymbol{\xi}_j)$ to the true measure $S(z)$.

Before affording the problem of estimating $S(z)$ from the data $\underline{\mathbf{a}}$ we need to check that the data provide enough information to solve it. Precise conditions that must be met to solve the problem are well known in the noiseless case and are reported in the introduction. When noise is present the identifiability problem is an open one. Its solvability can depend on the amount of "a priori" information available [6] and/or on the ability to devise smart algorithms. In the following a definition of identifiability is given and, based on it, some properties of $\mathbf{S}_n(z)$ are proved.

Definition 1 *The measure $S(z)$ is identifiable from the data $\mathbf{a_k}, k = 0, \ldots, n-1$ if $\exists\ r_k > 0, k = 1, \ldots, p$ such that*

- *$h_n(z)$ is unimodal in $N_k = \{z| \, |z - \xi_k| \leq r_k\}$*
- *$\bigcap_{k=1}^{p} N_k = \emptyset$*

The idea is that $S(z)$ can be identified from the data $\underline{\mathbf{a}}$ if the random generalized eigenvalues have a condensed density with separate peaks centered on $\xi_j, j = 1, \ldots, p$. As, by Theorem 2, $h_n(z, \sigma)$ converges weakly to $\frac{2}{n} \sum_{j=1}^{p} \delta(z - \xi_j)$ when $\sigma \to 0$, it must exists a $\sigma' > 0$ small enough to make $S(z)$ identifiable $\forall\ \sigma < \sigma'$.

In order to apply the proposed method one should check that the identifiability conditions are verified. As $h_n(z, \sigma)$ depends on the unknown quantities p, c_j, ξ_j this is of course impossible. However in most real problems we have some prior information about the unknown measure $S(z)$ that we can exploit to get reasonable interval estimates for p, c_j, ξ_j. Moreover in many instances either n or σ or both can be freely chosen. By Theorem 3, equation 11, n should not be as large as possible to get the best estimates of $S(z)$. In fact too many data will convey too much noise

which could mask the signal s_k. We can therefore properly design an experiment by computing $h_n(z,\sigma)$ for many values of n and σ and choose n_{ott} and σ_{ott} (optimal design) that make identifiable the measures corresponding to prior estimates of p, c_j, ξ_j. To identify the unknown measure $S(z)$ we then hopefully need to measure n_{ott} data affected by an error with s.d. σ_{ott}. Unfortunately $h_n(z)$ does not admit a closed form expression and to compute the expectation that appears in its definition we need to perform a time consuming MonteCarlo experiment. This is why we proposed an approximation $\tilde{h}_n(z)$ of $h_n(z)$ which can be quickly computed by solving hermitian eigenvalues problems.

Let us consider the function

$$S_n(z) = E[\mathbf{S}_n(z)] = \sum_{j=1}^{n/2} E[\mathbf{c}_j \delta(z - \boldsymbol{\xi}_j]$$

where $\{\mathbf{c}_j, \boldsymbol{\xi}_j\},\ j = 1, \ldots, n/2\}$ are the solution of the complex exponential interpolation problem for the data $\{\mathbf{a_k}, k = 0, \ldots, n-1\}$.

The relation between $S_n(z)$ and the unknown measure $S(z)$ is given by the following

Theorem 5 *If $S(z)$ is identifiable from $\underline{\mathbf{a}}$ then*

$$\int_{N_h} S_n(z)dz = c_h + o(\sigma)$$

and

$$\int_A S_n(z)dz = o(\sigma), \quad \forall A \subset D - \bigcup_j N_j$$

<u>proof</u>

From the identifiability hypothesis we know that

$$\int\limits_{N_k} h_n(z)dz = \frac{2}{n}\sum_{j=1}^{n/2} Prob[\boldsymbol{\xi}_j \in N_k] > 0, \quad k = 1,\ldots,p.$$

Therefore there exist $\boldsymbol{\xi}_{j_k}$ such that $Prob[\boldsymbol{\xi}_{j_k} \in N_k] > 0$. Among the $\boldsymbol{\xi}_{j_k}$ let us denote by $\boldsymbol{\xi}_{\hat{k}}$ the one such that $Prob[\boldsymbol{\xi}_{j_k} \in N_k]$ is maximum. From the identifiability hypothesis the $\boldsymbol{\xi}_{\hat{k}}$ are distinct. Moreover all the $\boldsymbol{\xi}_j,\ j = 1,\ldots,n/2$ can be sorted in such a way that $\boldsymbol{\xi}_j = \boldsymbol{\xi}_{\hat{j}},\ j = 1,\ldots,p$ and, by Lemma 2, to $\boldsymbol{\xi}_k$ it corresponds $\mathbf{c}_k$ such that

$$E[\mathbf{c}_k] = \begin{cases} c_k + o(\sigma),\, k = 1,\ldots,p \\ \\ o(\sigma), \qquad\quad k = p+1,\ldots,n/2 \end{cases}$$

But then for $k = 1,\ldots,p$

$$\int\limits_{N_k} S_n(z)dz = \sum_{j=1}^{n/2}\int\limits_{N_k} E[\mathbf{c}_j\delta(z-\boldsymbol{\xi}_j]dz =$$

$$= \sum_{j=1}^{n/2}\int\limits_{N_k}\left(\int\limits_{\mathbb{C}^2}\gamma\delta(z-\zeta)d\mu_{\gamma\zeta}\right)dz =$$

$$= \sum_{j=1}^{n/2}\int\limits_{\mathbb{C}^2}\gamma\left(\int\limits_{N_k}\delta(z-\zeta)dz\right)d\mu_{\gamma\zeta}$$

where $\mu_{\gamma\zeta}$ is the joint distribution of $\mathbf{c}_j$ and $\boldsymbol{\xi}_j$. We have

$$\int\limits_{N_k}\delta(z-\zeta)dz = \begin{cases} 1 \text{ if } \zeta \in N_k \\ \\ 0 \text{ otherwise} \end{cases}$$

hence,

$$\int\limits_{N_k} S_n(z)dz = \sum_{j=1}^{n/2} E[\mathbf{c}_j\delta_{jk}] = E[\mathbf{c}_k] = c_k + o(\sigma).$$

By a similar argument the second part of the thesis follows. $\square$

3 The P-transform

In order to solve the original moment problem we need to compute

$$S_n(z,\sigma^2)=\sum_{j=1}^{n/2}E[\mathbf{c}_j\delta(z-\boldsymbol{\xi}_j].$$

In order to estimate the expected value we build independent replications of the data (pseudosamples) by defining

$$\mathbf{a}_k^{(r)}=\mathbf{a}_k+\boldsymbol{\nu}_k^{(r)},\quad k=0,\ldots,n-1;\quad r=1,\ldots,R$$

where $\{\boldsymbol{\nu}_k^{(r)}\}$ are i.i.d. zero mean complex Gaussian variables with variance σ'^2 independent of $\mathbf{a}_h$, $\forall h$. Therefore

$$E[\mathbf{a}_k^{(r)}]=s_k,\quad E[(\mathbf{a}_k^{(r)}-s_k)(\overline{\mathbf{a}}_h^{(s)}-\overline{s}_h)]=\tilde{\sigma}^2\delta_{hk}\delta_{rs}$$

where $\tilde{\sigma}^2=\sigma^2+\sigma'^2$. For $r=1,\ldots,R$, we define the statistics

$$\hat{\mathbf{S}}_{n,r}(z,\tilde{\sigma}^2)=\sum_{j=1}^{n/2}\mathbf{c}_j^{(r)}\delta(z-\boldsymbol{\xi}_j^{(r)})$$

where $\mathbf{c}_j^{(r)},\boldsymbol{\xi}_j^{(r)}$ are the solution of the complex exponentials interpolation problem for the data $\mathbf{a}_k^{(r)}$, $k=0,\ldots,n-1$. As, by Lemma 2, the transformation

$$T:\{\mathbf{a}_k^{(r)},k=0,\ldots,n-1\}\rightarrow\{[\mathbf{c}_j^{(r)},\boldsymbol{\xi}_j^{(r)}],j=1,\ldots,n/2\}$$

is one-to-one, $\hat{\mathbf{S}}_{n,r}(z,\tilde{\sigma}^2)$ are i.i.d. with mean $S_n(z,\tilde{\sigma}^2)$ and finite variance $\zeta(z,\tilde{\sigma}^2)$ because $\{\boldsymbol{\nu}_k^{(r)}\}$ are i.i.d. . Therefore the statistic

$$\hat{\mathbf{S}}_{n,R}(z,\tilde{\sigma}^2)=\frac{1}{R}\sum_{r=1}^{R}\hat{\mathbf{S}}_{n,r}(z,\tilde{\sigma}^2)$$

has mean $S_n(z,\tilde{\sigma}^2)=E[\hat{\mathbf{S}}_{n,r}(z,\tilde{\sigma}^2)]$ and variance $\frac{1}{R}\zeta(z,\tilde{\sigma}^2)$.

Let us consider the statistic

$$\hat{\mathbf{S}}_n(z,\sigma^2) = \sum_{j=1}^{n/2} \mathbf{c}_j \delta(z - \boldsymbol{\xi}_j)$$

where $\mathbf{c}_j, \boldsymbol{\xi}_j$ are the solution of the complex exponentials interpolation problem for the data $\mathbf{a}_k, \quad k = 0, \ldots, n-1$ and the conditioned statistic

$$\hat{\mathbf{S}}^c_{n,R}(z,\tilde{\sigma}^2) = \hat{\mathbf{S}}_{n,R}(z,\tilde{\sigma}^2)|\underline{\mathbf{a}}$$

which are both computable from the observed data $\underline{a}$. We have

Lemma 3 *For n and $\sigma > 0$ fixed and $\forall z$ and $\tilde{\sigma}$,*

$$E[\hat{\mathbf{S}}^c_{n,R}(z,\tilde{\sigma}^2)] = S_n(z,\tilde{\sigma}^2)$$

$$\lim_{R\to\infty} var[\hat{\mathbf{S}}^c_{n,R}(z,\tilde{\sigma}^2)] = 0.$$

<u>proof</u>

from the conditional variance formula ([23]) we have

$$E[\hat{\mathbf{S}}^c_{n,R}(z,\tilde{\sigma}^2)] = E[\hat{\mathbf{S}}_{n,R}(z,\tilde{\sigma}^2)] = S_n(z,\tilde{\sigma}^2)$$

and

$$var[(\hat{\mathbf{S}}^c_{n,R}(z,\tilde{\sigma}^2)] \leq var[\hat{\mathbf{S}}_{n,R}(z,\tilde{\sigma}^2)] = \frac{1}{R}\zeta(z,\tilde{\sigma}^2). \qquad \square$$

It follows that $\forall z$ the risk of $\hat{\mathbf{S}}^c_{n,R}(z,\tilde{\sigma}^2)$ as an estimator of $S(z)$ with respect to the loss function given by the absolute difference could be smaller than the risk of the estimator $\hat{\mathbf{S}}_n(z,\sigma^2)$ if R and $\tilde{\sigma}$ are suitably chosen, despite of the fact that its bias is larger because $\tilde{\sigma} > \sigma$ and Theorem 5 holds. As a matter of fact this possibility is

always verified provided that σ' and R are suitably chosen as proved in the following

Theorem 6 *Let $M(z)$ and $M_c(z)$ be the mean squared error of $\hat{\mathbf{S}}_n(z, \sigma^2)$ and $\hat{\mathbf{S}}^c_{n,R}(z, \tilde{\sigma}^2)$ respectively. In the limit for $\sigma \to 0$, it exist σ' and $R(\sigma')$ such that $\forall R \geq R(\sigma')$, $M_c(z) < M(z) \ \ \forall z$.*

<u>proof</u>

let $M_c(z) = v_c + b_c^2$ and $M(z) = v + b^2$ be the decomposition of the mean squared errors in the sum of variance plus squared bias. Then $M_c(z) - b^2 = v_c + (b_c^2 - b^2)$. By Lemma 3, b_c is equal to the bias of $\hat{\mathbf{S}}_n(z, \tilde{\sigma}^2)$ and, by Theorem 5, it is $o(\tilde{\sigma})$ for $\tilde{\sigma} \to 0$. Then $\lim_{\sigma' \to 0^+}(b_c^2 - b^2) = 0$. Moreover, by Lemma 3, $\lim_{R\to\infty} v_c = 0$. Therefore $\forall v > 0, \ \ \exists \sigma'_v$ and $R(\sigma'_v)$ such that $\forall \sigma' < \sigma'_v, \ \ v_c + (b_c^2 - b^2) < v$ and then $M_c(z) < M(z)$. $\square$

In order to define a discrete transform, we evaluate $\hat{\mathbf{S}}^c_{n,R}(z, \tilde{\sigma}^2))$ on a lattice $L = \{(x_i, y_i), i = 1, \ldots, N\}$ such that

$$\min_j \Re\xi_j > \min_i x_i; \quad \max_j \Re\xi_j < \max_i x_i$$

$$\min_j \Im\xi_j > \min_i y_i; \quad \max_j \Im\xi_j < \max_i y_i.$$

In order to cope with the Dirac distribution appearing in the definition of $\hat{\mathbf{S}}^c_{n,R}(z, \tilde{\sigma}^2))$ it is convenient to use an alternative expression given by

$$\hat{\mathbf{S}}^c_{n,R}(z, \tilde{\sigma}^2) = \frac{1}{2\pi R} \Delta \left(\sum_{r=1}^{R} \sum_{j=1}^{n/2} [\mathbf{c}_j^{(r)} | \underline{\mathbf{a}}] \log(|z - [\boldsymbol{\xi}_j^{(r)} | \underline{\mathbf{a}}]|) \right)$$

which can be obtained by the former one by remembering that $\frac{1}{4\pi} \Delta \log(|z|^2) = \delta(z)$ (see e.g. [25, pg.47]). In this way the problem of discretizing the Dirac δ is reduced

to discretizing the Laplacian operator, which is easier to cope with. We then get a random matrix $\mathsf{P}(\tilde{\sigma}^2) \in \Re_+^{(N\times N)}$ such that $\mathsf{P}(h,k,\tilde{\sigma}^2) = \hat{\mathbf{S}}^c_{n,R}(x_h + iy_k)$. We call this matrix the P-transform of the vector $[\mathbf{a}_0, \ldots, \mathbf{a}_{n-1}]$.

4 Estimation procedure

The P-transform gives a global picture of the measure $S(z)$. However an estimate of the unknown parameters p, $\{\xi_j, c_j, j = 1, \ldots, p\}$ are usually of interest. An automatic procedure to get such estimates is now described. Let $\mathsf{P}(\tilde{\sigma}^2)$ be the P-transform computed by using R pseudosamples with variance $\tilde{\sigma}^2$. The proposed procedure is the following (dropping for simplicity the conditioning to $\underline{\mathbf{a}}$):

- memorize all the Pade' poles $\boldsymbol{\xi}_j^{(r)}$ and the corresponding residuals $\mathbf{c}_j^{(r)}, \quad r = 1, \ldots, R$ used for computing $\mathsf{P}(\tilde{\sigma}^2)$
- identify the local maxima of $\mathsf{P}(\tilde{\sigma}^2)$ and sort them in increasing order with respect to the local maxima values. The local maxima are candidate estimates of $\{\xi_j, j = 1, \ldots, p\}$
- for each candidate a cluster of (previously memorized) Pade' poles was estimated by including all the poles closest to the current candidate until the cluster cardinality equals a predefined percentage (e.g.$>$ 50%) of the number R of pseudosamples. The rationale is that if the candidate is close to one of the ξ_j most of the pseudosamples should provide a Pade' pole close to it. Notice that spurious clusters - i.e. not centered close to some ξ_j - can be expected [3]
- all the candidates whose associated cluster does not have the prescribed cardinality are eliminated. The number $\hat{p}$ of left candidates is then an estimate of

p

- for each of the $\hat{p}$ clusters the Pade' poles and the corresponding residuals (previously memorized) were then averaged and provided estimates $\hat{\xi}_j, \hat{c}_j, j = 1, \ldots, \hat{p}$ of the unknown parameters. Hopefully to $\hat{\xi}_j$ associated to spurious clusters should correspond relatively small $\hat{c}_j$.

5 Numerical results

In this section some experimental evidence of the claims made in the previous sections is given. A model with $p = 5$ components given by

$$\underline{\xi} = \left[e^{-0.1-i2\pi 0.3}, e^{-0.05-i2\pi 0.28}, e^{-0.0001+i2\pi 0.2}, e^{-0.0001+i2\pi 0.21}, e^{-0.3-i2\pi 0.35}\right]$$

$$\underline{c} = [6, 3, 1, 1, 20], \quad \sigma = 0.2, \quad n = 80$$

is considered. We notice that $SNR = 5$ and the frequencies of the 3^{rd} and 4^{th} components are closer than the Nyquist frequency $(0.21 - 0.20 = 0.01 < 1/n = 0.0125)$. Hence a superesolution problem is involved in this case. The quality of the approximation of $\tilde{h}(z)$ to the condensed density is first addressed, $\tilde{h}(z)$ is then computed along a line which pass through ξ_j and the closest among the $(\xi_h, h \neq j)$. If the model is identifiable $\tilde{h}(z)$ should have a local maximum close to ξ_j along this line. The interquartile range $\hat{r}_j$ of a restriction of $\tilde{h}(z)$ to a neighbor of this maximum is then considered as an estimate of the radius of the local support of $\tilde{h}(z)$ assumed circular. Then $M = 100$ independent data sets $\underline{a}^{(m)}$ of length n were generated and the Pade' poles $\underline{\xi}^{(m)}, m = 1, \ldots, M$ were plotted in fig.1 where circles of radii $\hat{r}_j$ centered on ξ_j have been represented too. We notice that the circles are reasonable

estimates of the Pade' poles clusters which provide an estimate of the support of the peaks of the true condensed density corresponding to $\xi_j, j = 1, \ldots, p$. We conclude that $\tilde{h}(z)$ is a reliable approximation of the condensed density and therefore, with the choice of n and σ made above, the model is likely to be identifiable.

We want now to show by means of a small simulation study the quality of the estimates of the parameters $\underline{\xi}$ and $\underline{c}$ which define the unknown measure $S(z)$. To this aim the bias, variance and mean squared error (MSE) of each parameter separately will be estimated. $M = 500$ independent data sets $\underline{a}^{(m)}$ of length n were generated by using the model parameters given above. For $m = 1, \ldots, M$ the P-transform $\mathsf{P}^{(m)}$ was computed based on $R = 100$ pseudosamples with $\sigma'^2 = 10^{-4}\sigma^2$ on a square grid of dimension $N = 200$. The estimation procedure is then applied to each of the $\mathsf{P}^{(m)}, m = 1, \ldots, M$ and the corresponding estimates $\hat{\xi}_j^{(m)}, \hat{c}_j^{(m)}, j = 1, \ldots, \hat{p}^{(m)}$ of the unknown parameters were obtained. As we know the true value p, if less than p local maxima were found in the second step or if $\hat{p}^{(m)} < p$ in the fourth step of the procedure, the corresponding data set $\underline{a}^{(m)}$ was discarded.

In Table 1 the bias, variance and MSE of each parameter including p is reported. They were computed by choosing among the $\hat{\xi}_j^{(m)}, j = 1, \ldots, \hat{p}^{(m)}$ the one closest to each $\xi_k, k = 1, \ldots, p$ and the corresponding $\hat{c}_j^{(m)}$. If more than one ξ_k is estimated by the same $\hat{\xi}_j^{(m)}$ the $m-$th data set $\underline{a}^{(m)}$ was discarded. In the case considered 65% data sets were accepted. Looking at Table 1 we can conclude that the true measure can be estimated quite accurately in 65% of cases.

When $\hat{p}_j^{(m)} > p$ we computed also the average residual amplitude

$$a_{res} = \frac{1}{|\tilde{M}|} \sum_{m \in \tilde{M}} \frac{1}{(\hat{p}^{(m)} - p)} \sum_{j=p+1}^{\hat{p}^{(m)}} \hat{c}_j^{(m)}, \text{ where } \tilde{M} = \{m | \hat{p}_j^{(m)} > p\}$$

which represents the contribution to $\hat{\mathbf{S}}^c_{n,R}(z, \tilde{\sigma}^2))$ of all the components which give rise to spurious clusters. In the case considered its value is $a_{res} = 1.165$ which should be compared with the true amplitudes $\underline{c}$. We can conclude that even when more components then the true ones are detected their relative importance is very low.

In order to appreciate the advantage of the estimator $\hat{\mathbf{S}}^c_{n,R}(z, \tilde{\sigma}^2)$ with respect to $\hat{\mathbf{S}}_n(z, \sigma^2)$, the same $M = 100$ independent data sets $\underline{a}^{(m)}$ of length n generated before were considered. The corresponding Pade' poles and weights $(\hat{\xi}_j^{(m)}, \hat{c}_j^{(m)}, j = 1, \ldots, n/2)$ were computed and ordered for each m in decreasing order w.r. to the absolute value of the weights. The true $(\xi_j, c_j, j = 1, \ldots, p)$ were ordered in the same way and the error

$$e_0(m) = \sum_{j=1}^{p} (\hat{\xi}_j^{(m)} - \xi_j)^2 + \sum_{j=1}^{p} (\hat{c}_j^{(m)} - c_j)^2$$

was computed for $m = 1, \ldots, M$ and plotted in fig.2. Then to each of the M data sets $\underline{a}^{(m)}$ previously generated $R = 100$ i.i.d. zero-mean Gaussian samples with variance $\sigma'^2 = 0.64\sigma^2$ were added and $(\hat{\xi}_j^{(m,r)}, \hat{c}_j^{(m,r)}, j = 1, \ldots, n/2, \quad r = 1, \ldots, R)$ were computed and ordered as before for each m and r. Finally the error

$$e_R(m) = \sum_{j=1}^{p} \left(\frac{1}{R} \sum_{r=1}^{R} \hat{\xi}_j^{(m,r)} - \xi_j \right)^2 + \sum_{j=1}^{p} \left(\frac{1}{R} \sum_{r=1}^{R} \hat{c}_j^{(m,r)} - c_j \right)^2$$

was computed for $m = 1, \ldots, M$ and plotted in fig.2. We notice that $e_R(m) \ll e_0(m)$ for almost all m and it is much less dispersed around its mean. Therefore the

estimates of $(\xi_j, c_j, j = 1, \ldots, p)$ obtained by averaging over the R pseudosamples are better than those obtained by the original samples. Finally we notice that in this simulation we used a variance $\tilde{\sigma}^2$ much larger than the one used to produce the results in Table 1. This large value gives the best mean squared error over all the five parameters but not necessarily the best reconstruction of each single parameter, as we looked for in the previous simulation.

References

[1] Barone, P. (2005). On the distribution of poles of Pade' approximants to the Z-transform of complex Gaussian white noise, *J. Approx. Theory* **132** 224-240.

[2] Barone, P., March, R. (2001). A novel class of Padé based method in spectral analysis. *J. Comput. Methods Sci. Eng.* **1** 185-211.

[3] Barone, P., March, R. (1998). Some properties of the asymptotic location of poles of Padé approximants to noisy rational functions, relevant for modal analysis. *IEEE Trans. Signal Process.* **46** 2448-2457.

[4] Barone, P., Ramponi, A.(2000). A new estimation method in modal analysis. *IEEE Trans. Signal Process.* **48** 1002-1014.

[5] Bharucha-Reid A.T., Sambandham M.(1986) *Random Polynomials.* Academic Press, New York.

[6] Donoho, D.L.(1992). Superresolution via sparsity constraints. *SIAM J. Math. Anal.*, **23,5** 1309-1331.

[7] Hammersley, J.M.(1956). The zeros of a random polynomial. *Proc. Berkely Symp. Math. Stat. Probability, 3rd*, **2** 89-111.

[8] Farrar, T.C. (1987) *Introduction To Pulse NMR Spectroscopy*, Farragut Press, Chicago.

[9] Flowe, R.P., Harris, G.A. (1993). A note on generalized Vandermonde determinants. *SIAM J. Matrix Anal. Appl.*, **14,4** 1146-1151.

[10] Gammel, J.L. (1972). Effect of random errors (noise) in the terms of a power series on the convergence of the Padé approximants. in *Padé Approximants*, Graves-Morris, P.R. ed., The Institute of Physics, London and Bristol.

[11] Gammel, J.L., Nuttall, J. (1973). Convergence of Padé approximants to quasianalytic functions beyond natural boundaries. *J. Math. Anal. Appl.* **43** 694-696.

[12] Golub, G.H., Milanfar, P., Varah, J.(2004). A stable numerical method for inverting shapes from moments. *SIAM J. Sci. Comp.*,bf 21,4 1222–1243.

[13] Green, R.E., Krantz, S.G.(1997). *Function theory of one complex variable*, John Wiley, New York.

[14] Henrici, P.(1977). *Applied and computational complex analysis, vol.I*, John Wiley, New York.

[15] Jialiang, Li (1993). On the existence and convergence of random Padé approximants, *Adv. Math.* **22** 340-347.

[16] Kravanjaa, P., Sakuraib, T., Sugiurac, H., Van Barel, M. (2003). A perturbation result for generalized eigenvalue problems and its application to error estimation in a quadrature method for computing zeros of analytic functions. *J. Comput. Appl. Math.* **161** 339-347.

[17] Krattenthaler, C. (1998) Advanced determinant calculus. *Séminaire Lotharingien Combin. 42 ("The Andrews Festschrift")*, **Article B42q**.

[18] Lehmann, E.L.(1983). *Theory of point estimation*, Wiley, New York.

[19] March, R., Barone, P.(1998). Application of the Padé method to solve the noisy trigonometric moment problem: some initial results. *SIAM J. Appl. Math.* **58** 324-343.

[20] March, R., Barone, P.(2000). Reconstruction of a piecewise constant function from noisy Fourier coefficients by Padé method. *SIAM J. Appl. Math.* **60** 1137-1156.

[21] Markushevich, A.I.(1965). *Theory of functio ns of a complex variable vol.II*, Prentice-Hall, Englewood Cliffs, N.J.

[22] Osborne M.R., Smyth G.K.(1995), A Modified Prony Algorithm for Exponential Function Fitting,*SIAM J. Sci. Comput.* **16** 119-138.

[23] Renyi, A.(1970). *Probability Theory*, North Holland Pub. Co., Amsterdam.

[24] Scharf, L.L.(1991). *Statistical signal processing*, Addison-Wesley, Reading.

[25] Schwartz, L.(1950). *Théorie des distributions, vol.1*, Hermann,Paris.

[26] Stewart, G.W.(2001). *Matrix algorithms, vol.2*, SIAM, Philadelphia.

[27] Stoer, J., Bulirsch, R.(1996). *Introduction to numerical analysis*, Springer-Verlag, New York.

[28] Viti, V., Petrucci, C. and Barone, P.(1997). Prony methods in NMR spectroscopy. *International Journal of Imaging Systems and Technology* **8** 565-571.

	p	$bias(\hat{p})$	$s.d.(\hat{p})$	$MSE(\hat{p})$
	5	0.0500	1.0000	1.0025
	ξ_j	$bias(\hat{\xi}_j)$	$s.d.\hat{\xi}_j$	$MSE(\hat{\xi}_j)$
$j=1$	-0.2796 - 0.8606i	-0.0006 + 0.0004i	0.0230	0.0005
$j=2$	-0.1782 - 0.9344i	-0.0005 - 0.0004i	0.0125	0.0002
$j=3$	0.3090 + 0.9510i	0.0057 - 0.0009i	0.0171	0.0003
$j=4$	0.2487 + 0.9685i	-0.0005 + 0.0024i	0.0145	0.0002
$j=5$	-0.4354 + 0.5993i	-0.0054 + 0.0018i	0.0290	0.0009
	c_j	$bias(\hat{c}_j)$	$s.d.(\hat{c}_j)$	$MSE(\hat{c}_j)$
$j=1$	6.0000	0.1545	1.7154	2.9663
$j=2$	3.0000	-0.1617	1.2865	1.6812
$j=3$	1.0000	-0.1037	0.3295	0.1193
$j=4$	1.0000	-0.0981	0.3193	0.1116
$j=5$	20.0000	-0.1759	2.5101	6.3317

Table 1

Statistics of the parameters $\hat{p}$, $\hat{\xi}_j, j=1,\ldots,p$ and $\hat{c}_j, j=1,\ldots,p$

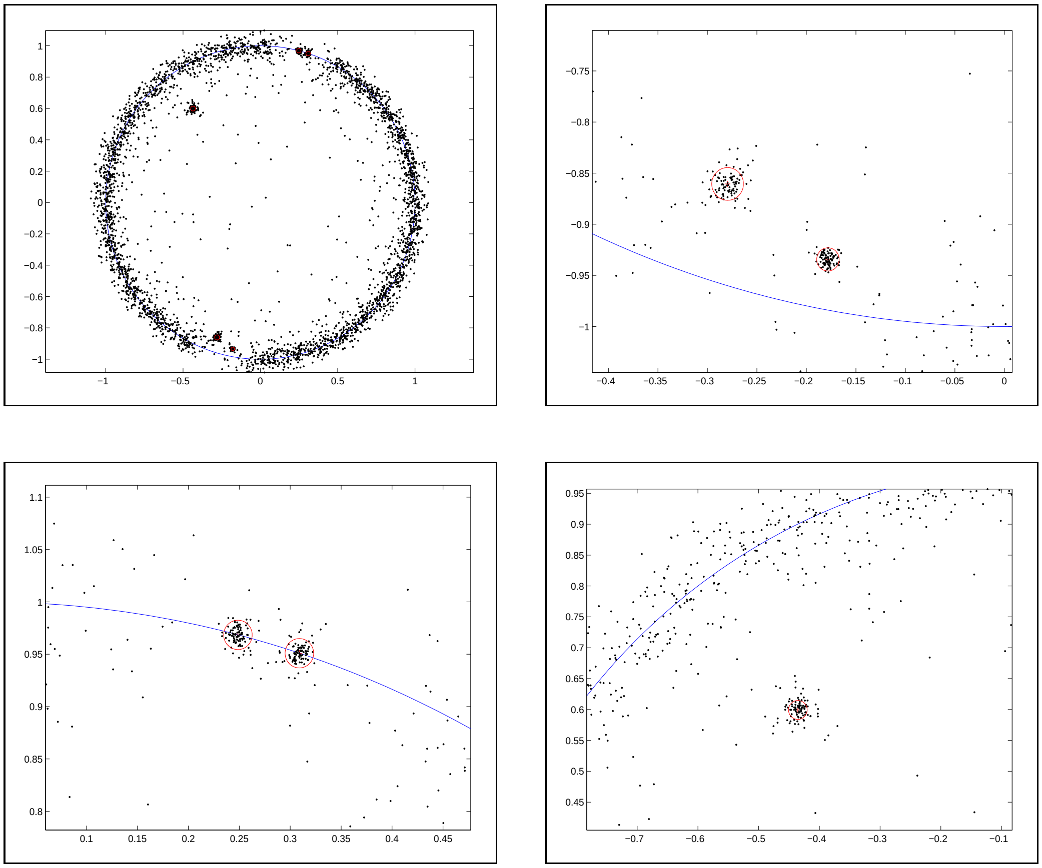

Fig. 1. Top left: location of Pade' poles for 100 independent realizations of the noise; the circles are the estimated support of the condensed density in a neighborhood of ξ_j; top right:zoom in a neighborhood of the 1-st and 2-nd components; bottom left: zoom in a neighborhood of the 3-rd and 4-th components; zoom in a neighborhood of the 5-th component (see section 4).

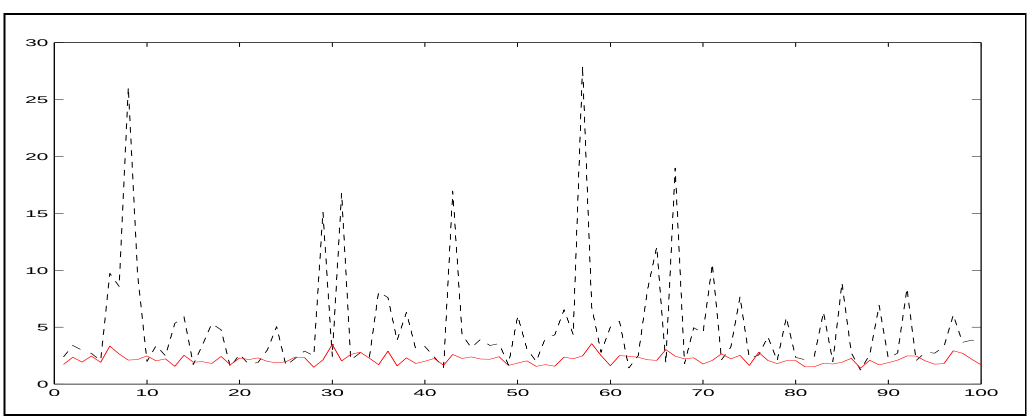

Fig. 2. MSE of the standard estimator of the parameters $(\xi_j, c_j), j = 1, \ldots, p$ (dashed); MSE of the averaged estimator (solid)